IMAGE RECOGNITION FOR VISUALLY IMPAIRED PEOPLE BY SOUND

K.GOPALA KRISHNAN, C.M.PORKODI, and K.KANIMOZHI

Abstract— **The aim of this paper is to represent a method where a blind person can get information about the shape of an image through speech signal. In this paper we proposed an algorithm for image recognition by speech sound. Blind people face a number of challenges when interacting with their environments because so much information is encoded visually. The proposed method enables the visually impaired people to see with the help of ears. The novelty of this paper is to convert the image to sound using the methodology of edge detection.**

Index Terms— Edge detection, Feature Extraction, Image classification, Sensory replacement, Sound generation.

INTRODUCTION

Visual impairment is the major disability faced by millions of people around the world. Humans are very sensitive to sound. They employ multiple channels (sounds, vision, hap-tic etc) in their everyday lives to support many of their day-to-day routines. Humans easily identify a wide variety of sounds such as splashes, pouring, streaming, breaking waves and dripping. Exploiting this human capability we can convey many kinds of information to the visually impaired. It is possible to create an impression of objects or things that a person sounds and selected suitable musical parameters. If the sound scape of events can be conveyed via the auditory channel, it certainly helps the task of understanding events, objects and textures. It is therefore possible to substitute the eyes with ears to represent the physical world to some extent. This paper is divided into nine parts. The first part of paper deals with related work, second part deals with the methodology and proposed algorithm, and the last part is conclusion.

I. LITERATURE

Many interesting proposals have been done in the area of image recognition for people with visual disabilities. In 1940-1950 a device was developed by Dr. Cooper at Haskins Laboratory which was used to play a tone corresponding to a pattern and the tone could be used to recognize the pattern. Our idea inspired by the novel of human ability of perception. Sensory substitution method in which any sensory impairment can be replaced by another one. By this hearing can be used as the substitution for blindness. This paper addresses the technical feasibility of replacing human vision by human hearing through equipment that translates images into sounds.

II. SYSTEM OVERVIEW

The conceptual model is illustrated in Fig 1.

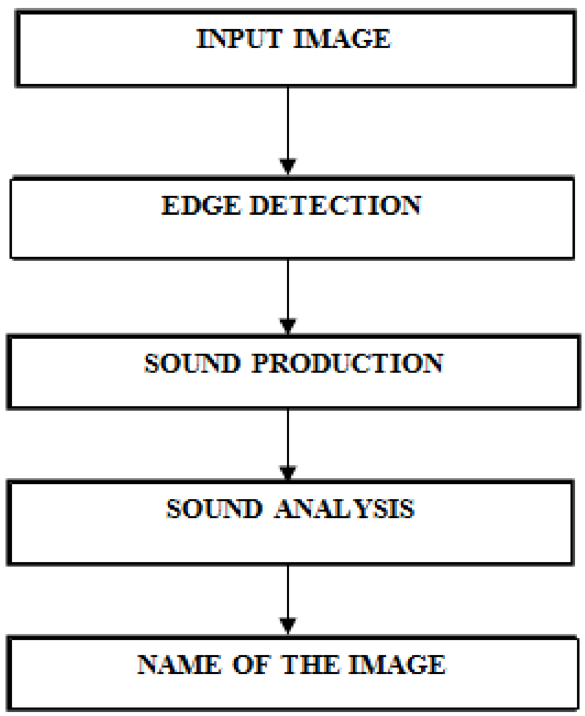


Fig 1 .System Overview

III. EDGE DETECTION

Edge detection is the first module and it is one of the pre-processing steps [4].Edge may be defined as a boundary between two homogeneous regions. Important features can be extracted from the edges of an image [5].In our project we use canny operator to detect the edges. An edge is typically extracted by computing the derivative of the image intensity function. This consists of two parts:

1. Magnitude of the derivative: measure of the Strength/contrast of the edge.

F. A. K.Gopala Krishnan is an Assistant Professor in the Mepco Schlenk Engineering College,Sivakasi.(e-mail:gopi1969@yahoo.com)
S. B. C.M.Porkodi is a student in Mepco Schlenk Engineering College,Sivakasi.(e-mail:porkodibeece@gmail.com)
T. C. K.Kanimozhi is a student in Mepco Schlenk Engineering College,Sivakasi.(e-mail:sweetkani25@gmail.com)

978-1-4673-4866-9/13/$31.00 ©2013 IEEE

2. Direction of the derivative vector: edge orientation.

IV. STEPS IN EDGE DETECTION

A. *Smoothing*

The image acquired by camera contains certain amount of noise. To prevent that noise is mistaken for edges, noise must be reduced. Therefore the image is first smoothed by applying a Gaussian filter. Mathematically, applying a Gaussian blur to an image is the same as convolving the image with a Gaussian function. Since the Fourier transform of a Gaussian is another Gaussian, applying a Gaussian blur has the effect of reducing the image's high-frequency components; a Gaussian blur is thus a low pass filter [6][7].The kernel of Gaussian filter with the standard deviation of σ= 1.4 is shown in Fig 2.

B=159

2	4	5	4	2
4	9	12	9	4
5	12	15	12	5
4	9	12	9	4
2	4	5	4	2

Fig 2. Gaussian filter

B. *Finding Gradients*

Edges can be determined by finding gradients. To find gradients we use sobel operator [4].The mask used for sobel operator is shown in Fig 3.

-1	0	+1
-2	0	+2
-1	0	+1

G_x

+1	+2	+1
0	0	0
-1	-2	-1

G_y

Fig 3. Mask of sobel operator

Where Gx is used to find gradient in horizontal direction and Gy is used find gradient in vertical direction. The magnitude of image gradient is calculated by using the law of Pythagoras as,

$$|G| = \sqrt{Gx^2 + Gy^2} \quad (1)$$

It is sometimes simplified by applying Manhattan distance measure as,

$$|G|=|Gx|+|Gy \quad (2)$$

The gradients can be calculated using these formulas [5]:

Gx [i, j] = (s[i, j+1]-s[i, j]+s[i+1, j+1]-s[i+1, j])=2 (3)

Gy [i, j] = (s[i, j]-s[i+1,j]+s[i, j+1]-s[i+1,j+1])=2 (4)

The gradient magnitudes often indicate the edges quite clearly. However, the edges are typically broad and thus do not indicate exactly where the edges are. To make it possible to determine this, the direction of the edges must be determined and stored as,

$$\theta = Gy/Gx \quad (5)$$

C. *Nonmaximum suppression*

After the edge directions are known, non-maximum suppression now has to be applied [6]. Non-maximum suppression is used to trace along the edge in the edge direction and suppress any pixel value (sets it equal to 0) that is not considered to be an edge. This will give a thin line in the output image. For example:

1. If the rounded angle is zero degrees the point will be considered to be on the edge if its intensity is greater than the intensities in the north and south directions.
2. If the rounded angle is 90 degrees the point will be considered to be on the edge if its intensity is greater than the intensities in the west and east directions.
3. If the rounded angle is 135 degrees the point will be considered to be on the edge if its intensity is greater than the intensities in the north east and south west directions.
4. If the rounded angle is 45 degrees the point will be considered to be on the edge if its intensity is greater than the intensities in the north west and south east directions.

D. *Double Thresholding*

. Canny edge detector uses two threshold values [8]. Edge pixels stronger than the high threshold are marked as Strong. Edge pixels weaker than the low threshold are suppressed and edge pixels between the two thresholds are marked as weak.

The result of edge detection is shown in Fig 4.

Fig 4. Edge detected image

V. SOUND GENERATION

In this module the sound is generated based on the edge information obtained from the steps above [1][9][10][11].The sound is generated using the MATLAB'S syntax,

Sound(y, Fs)

The sound hardware will convert the vector of numbers y into a sound waveform at a certain rate, called the sampling rate(Fs).Values in the y are assumed to be in the range of $-1.0<y<+1.0$. Values outside the range are clipped.

To differentiate the presence of horizontal and vertical edges in the image two different frequencies are used. Low frequency is used for horizontal edges and high frequency is used for vertical edges. Therefore frequency of the sound will be selected according to the orientation of the edge (Horizontal or vertical) and duration of sound will be decided by the length of the edge [1].The generated sound can also be plotted. It is shown in Fig .5

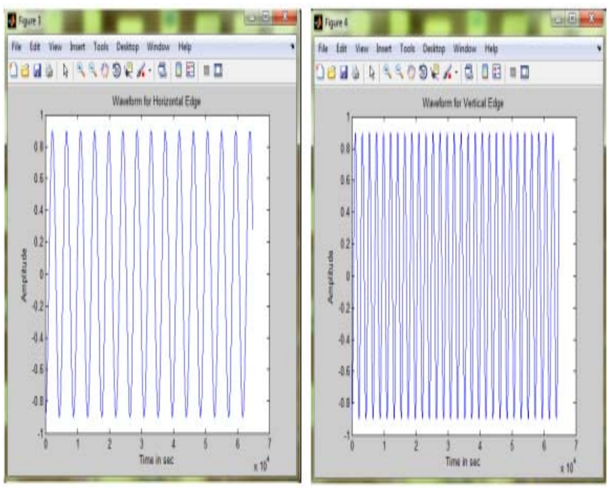


Fig 5. Waveform of sound

VI. SOUND ANALYSIS

The generated sound can be analyzed by finding the statistical properties like mean from the wavelet coefficients [12]. The statistical properties differ based on duration of sound (depends on length of the image).By these statistical properties the images can be classified as different categories. Using microphone the category of the image is recorded in human voice and the appropriate voice will be played according to the range of the statistical properties.

VII. IMPACT ON SOCIETY

After this the blind people can sit in front of the system and they can tell the name of the image by hearing the voice. By this work eyes can be replaced by ears [13][14].

VIII. RESULTS

From the work the result obtained by each module is listed below. In first module (Edge detection) canny operator is used to find the edges. The result of edge detection is shown in Fig 6 and Fig 7.

Fig 6.Edge detected image

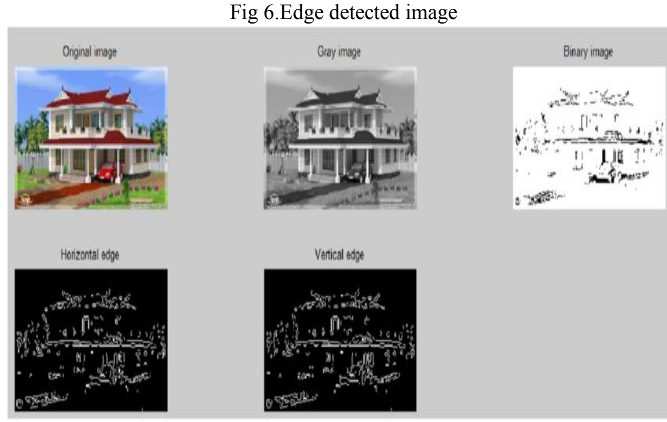
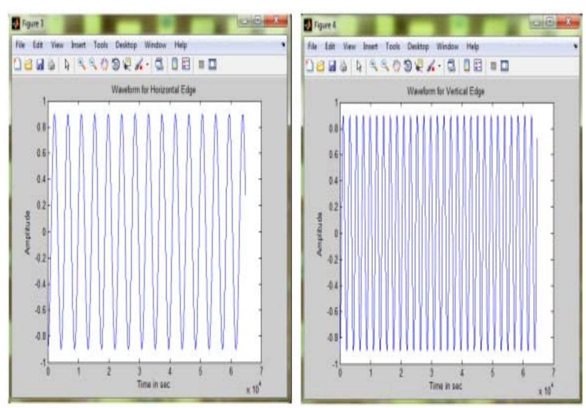


Fig 7.Edge detected image

In second module sound is produced using matlab. The Waveform of produced sound is plotted and is shown below. By this one can interpret the presence of low frequency (horizontal edge) and high frequency (vertical edge). The result of this module is shown in Fig 8.



(a) (b)

Fig 8. (a).Waveform of horizontal edge, (b).Waveform of vertical edge.

In third module the produced sound is analyzed using the statistical property like mean of wavelet coefficients. The result of this module is shown in Table 1

TABLE 1
RESULT OF SOUND ANALYSIS

IMAGE	RECOGNITION RATE
	10/15
	10/15
	8/15
	7/15
	2/15
	3/15

From the Table 1 it is intercepted that simple images like square,circle,cup,chair,rose,dog etc can be analyzed optimally and images like house, trees(contains more number of edges) etc requires more number of features to be considered to analyze the images.

IX. CONCLUSION

From the analysis, it is therefore possible to substitute the eyes with ears to represent the physical world to some extent. Our system is being presented as a solution for the conversion of images into sound. It is designed to be portable and low power. We can see the system works great for simple images, the performance decreases with the complexity in image. To improve the performance more number of features is to be considered. The development towards an efficient and practical application still needs to be finalized.

REFERENCES

[1] Zubair Khan, Saurabh Singh, Gaurav Agarwal,"Image Recognition by Sound Pattern Generated by the Image". International Journal of Computer Applications (0975 to 8887).Volume 7, No.12, October 2010.

[2] Manjur Alam,"Image Shape Approximation and Representing it by Speech Sound Pattern for Visually Impaired Person'. International Journal of Recent Trends in Engineering, Vol 1, No. 3, May 2009.

[3] Suresh Matta, Heiko Rudolph, and Dinesh K Kumar, "Auditory Eyes: Representing Visual Information In Sound And Tactile Cues", September, 2005.

[4] Mr.Manoj, K.Vairalkar, Prof.S.U.Nimbhorkar, "Edge Detection of Images Using Sobel Operator", International Journal of Emerging Technology and Advanced Engineering.(ISSN 2250 to 2459) Volume 2,Issue 1,January 2012.

[5] Andreas Koschan, Mongi Abidi,"Detection and Classification of Edges in Color Images".IEEE Signal Processing Magazine, January, 2005.

[6] Jasmeet Singh,"Canny algorithm for Edge Detection in Digital Images", 2000.

[7] Abravanel, E.Active detection of solid-shape Information by touch and vision. Perception and Pshycophysics, 10,358-360,, 1971.

[8] John Canny, "A Computational approach to Edge Detection". Pattern analysis and Machine Intelligence, IEEE Transactions on, PAMI-8(6),679-698,November,2010.

[9] Ehsan Nadernejad,Sara Sharifzadeh,Hamid Hassanpour,"Edge Detection Techniques: Evaluations and Comparisons".Applied Mathematical Sciences, Vol. 2,no. 31, 1507 to 1520,2008.

[10] Drd.Alexandru Cazan, prof. Radu Varbanescu and prof. Dan Popescu,"Algorithms and Techniques for Image to Sound Conversion for Helping the Visually Impaired People, Application Proposal", 2007.

[11] Peter B. L. Meijer, 'An Experimental System for Auditory Image Representations".IEEE Transactions ON Biomedical Engineering, Vol. 39, NO. 2, February 1992.

[12] T.Lumbrou, P.Kudumakis, R.Speller, M.Sandleg, A.Linney,"Classification of Audio Signals Using Statistical Features on Time and Wavelet Transform Domains".IEEE International Conference on speech and signal processing. Volume 6, 3621-3624, 1998.

[13] P.Arno, A.Vanlierde, E.Streel, M.C.Wanet Defalque, S.SanabriaBohorquez, C.Veraart,"Auditory Substitution of Vision: Pattern Recognition by the Blind".Applied cognitive Psychology.Appl.cognit.psychol, 15: 509 - 519, 2001.

[14] Bach-y-Rita, P."Sensory plasticity: Applications to a vision substitution system". Acta Neurologica Scandanavica, 43,417 to 426, 1967.

[15] Bach-y-Rita, P."Brain mechanisms in sensory substitution".New York: Academic Press. Bliss, J.C., M.H. Katcher, C.H. Rogers, and R.P. Shepard. 1970. Optical-to- tactile image, 1972.

[16] M. Capp and P. Picton,"The Optophone: an Electronic Blind Aid", The Engineering Science and Education Journal, 2000.

[17] L. Caponetti, N. Abbattists, and G. Carapella,"A genetic approach to edge detection", in Int. Conf. Image Processing, vol. 94, 1994, pp. 318 to 322.

[18] E. A. S. Galvanin, G. M. do Vale, and A. P. Dal Poz,"The Canny Detector with Edge Region Focusing Using an Anisotropic Diffusion Process".Sao Paulo State University, Department of Cartography, 305 Rua Roberta Simonsen, Presidente Prudente,Saeo Paolo, 19060-900 Brazil.

[19] M.I.Khalil,"Image Stegnography: Hiding short audio messages within Digital Images".JCS and T Vol. 11 No. 2 October 2011.

[20] H.B.Kekre Sr. Professor,Archana Athawale Asst.Prof., Thadomal,Swarnalata Rao,Uttara Athawale,MCA Dept.,"Information Hiding in Audio Signals''. International Journal of Computer Applications (0975 to 8887) Volume 7, No.9, October 2010.